

Native and Non-Native Patterns in Conflict: Lexicon vs. Grammar in Loanword Adaptation in Brazilian Portuguese

Natália Brambatti Guzzo
McGill University

1 Introduction

Research on loanword adaptation has shown that the process is affected by both perceptual and phonological factors, which require, respectively, faithfulness to the source and compliance with the phonotactic requirements of the borrowing language (e.g., Silverman (1992); Kang (2003, 2011); Davis & Cho (2006); Yip (2006)). Although some views on loanword adaptation argue for the primacy of either perceptual factors (see Peperkamp & Dupoux (2003); Peperkamp (2004); Peperkamp et al. (2008)) or phonological factors (see Paradis & LaCharité (1997); LaCharité & Paradis (2005)), both agree, at least to a certain extent, that speakers' perception of non-native sounds is affected by the phonology of the borrowing language.

In this paper, I discuss another possible factor that impacts loanword phonology: that the lexicon and the grammar behave differently with regard to faithfulness to the source and phonotactic violations in the adaptation of loanwords. It has been shown that the lexicon and the grammar can display inconsistencies, in that native speakers fail to generalize lexical patterns that are unnatural from a crosslinguistic perspective (Becker et al., 2012; Garcia, 2017; Jarosz, 2017). This paper extends evidence for this type of asymmetry to another domain, namely, loanword phonology. In this scenario, newly-introduced loanwords are produced in a way that is different from the way in which loanwords that are already part of the lexicon are produced. While loanwords in the lexicon are not constrained by the phonotactic requirements of the borrowing language, new loanwords are filtered by the phonology. Consequently, new loanwords, but not the loanwords existing in the lexicon, conform to the native patterns.

I probe this possibility by analyzing the adaptation patterns of English words containing /ʌ/ in Brazilian Portuguese (BP). As discussed below, the BP lexicon includes several loanwords containing /ʌ/, which is normally produced as [ɐ]. In the native BP inventory, [ɐ] is the closest sound to /ʌ/ (Barbosa & Albano, 2004), but it is severely constrained in native BP words: it is only allowed before a nasal consonant. The question that arises is whether native speakers extend the native contextual constraints on the use of [ɐ] to new loanwords, or whether they follow the generalization that is present in the lexicon (i.e., which allows [ɐ] in all loanwords with /ʌ/, regardless of the quality of the following segment).

To answer this question, I designed two production experiments, one with real loanwords and another one with nonce loanwords. Before I discuss the methodology and the results of these experiments (sections 3 and 4, respectively), I describe the requirements of BP phonology with regard to [ɐ] and how they impact the adaptation of English loanwords containing /ʌ/ (section 2). In section 5, I formalize the results of the Nonce Loanword Experiment using a probabilistic model of grammar (in line with, e.g., Goldwater & Johnson (2003); Hayes & Wilson (2008)). Section 6 concludes.

2 English loanwords with /ʌ/ in Brazilian Portuguese

In BP, the vowel that is articulatorily and acoustically closest to English /ʌ/ is [ɐ]. English loanwords containing /ʌ/ are typically produced with [ɐ] in BP, as shown in the examples in (1). Given that BP does not

* Many thanks to Heather Goad for her input on several aspects of this study. I also thank Guilherme D. Garcia, Michael Dow, Maren Hamilton, the native speakers of Brazilian Portuguese who participated in this study, and the audience at AMP 2018.

allow CVC syllables with an obstruent coda (other than /s/), we also observe the insertion of an epenthetic [ɪ] in the examples below.

- (1) a. *bug* → ['bɛgɪ], *pub* → ['pɛbɪ]
 b. *funk* → ['fẽᵑkɪ], *punk* → ['pẽᵑkɪ]

However, BP [ɐ] is an allophone for /a/ and appears only in nasal contexts (either in a syllable with a tautosyllabic nasal coda (2b) or in a stressed CV syllable followed by a nasal onset (2c)).

- (2) a. *casa* ['kaza], *['keza] 'house'
 b. *canto* ['kẽᵑtɔ], *['kaᵑtɔ], *['kãᵑtɔ] 'corner'
 c. *cama* ['kẽma], *['kama], *['kãma] 'house'

It should be noted that, in the literature on BP, it has been proposed that items such as (2b) exhibit *phonological* nasality, while items such as (2c) exhibit *phonetic* nasality (e.g., Leite (1974); see also Guimarães & Nevins (2013)). For analyses that follow this assumption, the crucial difference between the two forms of nasality is that, in phonological nasality, the vowel is nasal underlyingly (e.g., /ẽ/). On the other hand, it has also been argued that items such as (2b) have a nasal coda, and that the change in vowel quality in both (2b) and (2c) is due to feature spreading (see e.g., Câmara Jr. (1970); Battisti (1997); Wetzels (1997)). Since the vowel resulting from nasalization in (2b) and (2c) is articulatorily and acoustically the same (Cagliari, 1977), this paper does not intend to arbitrate between the two analyses. Additionally, word-final unstressed /a/ is usually transcribed as [ɐ] (Barbosa & Albano, 2004). In this position, vowels undergo reduction in BP. Since reduction in BP is a gradient process, speakers may alternate between [a] and [ɐ] (and forms in-between these two) in their productions (see e.g., Massini-Cagliari (1992)). However, in nasal contexts, as well as in stressed positions, alternation between [a] and [ɐ] is not an option in BP phonology, as indicated in the examples in (2).

Given the native BP inventory, while it is not surprising that [ɐ] is licensed in the items in (1b), it is surprising that such a vowel is licensed in the items in (1a), where there is no phonotactic context for [ɐ] to arise. In other words, if the adaptation of such loanwords had been strictly phonological, we should expect to see [a] in the items in (1a), and [ẽ] in the items in (1b).

It turns out that certain English loanwords with /ʌ/ have been adapted with [a] in BP, such as [pi'kapɪ] (from *pick-up*, meaning 'pick-up truck') and [ʃe'kapɪ] (from *check-up*, meaning 'medical check-up'). These examples suggest that at least some of the English loanwords that entered the BP lexicon were filtered by the native phonological grammar, which rejected [ɐ] in non-nasal contexts in favor of [a]. However, the adaptation of /ʌ/ into [a] in the BP lexicon seems to be the exception, not the rule. Another possibility of adaptation for loanwords with /ʌ/ is their production with [u] in BP (e.g., ['bugɪ] for *bug*). This, however, is an effect of orthography, as orthographic *u* is invariably produced as [u] in native Portuguese words. I return to this in the examination of the experimental results.

Given the examples in (1) and the native forms in (2), one question that must be addressed is the following: is the ability to license [ɐ] in more contexts in loanwords than in native words part of the BP grammar? Or, in other words: is this pattern of adaptation only present in the lexicon, but not in the grammar? If the lexicon and the grammar do differ with respect to the preferred adaptation patterns, then BP illustrates a case of lexicon-grammar asymmetry in the generalization of non-native patterns that are present in the lexicon due to borrowing.

Instances of lexicon-grammar asymmetries have been attested in native languages where unnatural lexical patterns failed to be generalized by native speakers (Becker et al., 2012; Garcia, 2017; Jarosz, 2017). In these situations, however, the lexicon was observed to contain patterns that, although unnatural from a crosslinguistic perspective, were not necessarily introduced through borrowing. When a potential asymmetry between the lexicon and the grammar arises in a loanword adaptation context, it is unclear whether speakers would prefer to adapt the new loanwords following the patterns of their native phonological grammars or the non-native patterns that are exhibited by the loanwords that are already part of the lexicon.

To address these issues, two experiments were conducted. The next section describes the methodology of these experiments.

3 Methodology

The production of English loanwords with /ʌ/ was tested through two experiments, one of which involved real loanwords and another one involving nonce loanwords. Native speakers of BP ($n = 15$) with various levels of proficiency in English were recruited. Participants' proficiency was estimated based on their responses to a questionnaire where they self-reported their language use and estimated their proficiency in English. The questionnaire included questions such as the amount of weekly exposure to English, years of contact with English, and years of English learning in a formal (school) setting. Participants also recorded a short passage, which was used to estimate their speaking proficiency in English. Participants' proficiency ranged from low intermediate to advanced; for the purposes of the present analysis, participants were divided into two groups (intermediate and advanced).

All of the participants were born in Brazil and had been living in Montreal for at most two years at the time of the experiment. Although the questionnaire included a question about participants' place of origin in Brazil, this variable was not controlled for in the study, since there are no dialectal differences in the distribution of [v] for the structures exemplified in (2).

All of the participants completed both experiments. Participants did the Real Loanword Experiment first, and, after a break of approximately ten minutes, did the Nonce Loanword Experiment. Both experiments were conducted in a sound-attenuated booth in the Department of Linguistics at McGill University. Participants were compensated for their time.

The following subsections describe each of the experiments.

3.1 Real Loanword Experiment The target items in this experiment were relatively frequent English loanwords with /ʌ/ ($n = 26$). The loanwords corresponded to brand, product or person names that are known by Brazilians (e.g., *Subway*, *Candy Crush*, *Trump*), and to nouns that are likely to be used in the media or everyday conversations (e.g., *pub*, *plug*, *funk*). Eleven of the target items had a nasal coda (e.g., *Trump*, *funk*). Loanwords such as *pick-up* and *check-up* were not included in the experiment; as these words are consistently adapted with an [a], they are optionally written in Portuguese with the letter *a* instead of *u* (e.g., *picape*, *checape*). For consistency, the target items used in this experiment are written with the letter *u*.¹

The loanwords were placed in newspaper headlines, which were presented to the participants on a computer screen. Participants saw one headline at a time and were instructed to read each headline as it appeared on the screen. After reading the headline, participants pressed a key and moved to a blank screen. While staring at the blank screen, participants were instructed to repeat the headline as faithfully as possible. With this procedure, the objective was to have two productions of each loanword by participant, one without the aid of an orthographic form. Specifically, I predicted that while their first productions could be constrained by the orthographic form of the loanword (which would result in more productions with [u]), their second productions would be more likely to exhibit [v]. However, at times some of the participants avoided producing the loanword in their second production, and either used a synonym or paraphrased the entire headline instead.

The total number of headlines in the Real Loanword Experiment was 35, 24 of which included the target loanwords (some of the headlines included more than one target loanword). The filler headlines contained either English loanwords of other shapes (e.g., *McDonalds*, *heavy metal*) or only native Portuguese words. Some of the target sentences also included additional loanwords of other shapes.

Participants took about 15 minutes to complete this experiment.

3.2 Nonce Loanword Experiment The target items in this experiment were nonce loanwords containing /ʌ/ ($n = 20$). All of the experimental items were monosyllabic (including the fillers; $n = 20$). Half of the target items had a nasal coda (e.g., *tump* [tʌmp], *vunt* [vʌnt]), while half of them had an oral coda (e.g., *vup* [vʌp], *cush* [kʌʃ]).

For this experiment, participants were told that they would be presented with words that referred to alien species that had been recently discovered by English-speaking scientists, and that their task was to produce these words in sentences in Portuguese. Participants saw the image of an alien on the computer screen and, after a beep, heard the alien name twice (the alien images were retrieved from the internet using Google's search tool). At this point, participants were instructed to say *This is a [loanword]* in Portuguese. The experiment then automatically moved to the next screen, where participants were presented with the

¹ The only exception is the loanword *country*.

picture of the favorite activity of the alien in question (only three possibilities were available: painting, sports, and music); see Figure 1. Participants were then instructed to say *The [loanword] likes [depicted activity]*. Participants took approximately 15 minutes to complete this experiment. The stimuli were recorded by a native speaker of American English with training in linguistics and phonetics.



Figure 1: Example of screen in the Nonce Loanword Experiment.

Participants were presented with two opportunities to produce each loanword for two reasons: (i) to minimize potential effects of imitation, and (ii) to examine the possibility that their second production would be more phonologically constrained than the first one. Specifically with respect to (ii), I predicted that participants' first production of the nonce loanwords could be primarily constrained by their perception of the item, while their second production could be adapted to the phonotactic requirements of their native language. If this were the case, for the nonce loanwords where /ʌ/ precedes an oral coda, we should expect to observe [ɐ] more frequently in their first productions. In their second productions, [a] should be observed more frequently, as it is the phoneme corresponding to [ɐ] in oral contexts in BP native phonology. No orthographic stimuli was provided in this experiment, so productions with [u] were not expected.

3.3 Predictions If the lexicon and the grammar shape loanwords with /ʌ/ differently, the overall prediction for the two experiments is that [ɐ] will be more frequently produced in the Real Loanword Experiment than in the Nonce Loanword Experiment, both in nasal and oral contexts. In other words, [ɐ] will be produced more frequently in items that are already part of the BP lexicon, regardless of the context where the target vowel appears. However, when speakers act as the disseminators² of new loanwords, they may filter them through their native phonological grammar. In this case, [ɐ] will be blocked in non-nasal contexts, and [a] will be used instead. The next section discusses the results of these two experiments in light of these predictions.

4 Data and results

4.1 Overall results For both experiments, participants' productions of the target vowel were segmented and coded for vowel quality and for whether they preceded an oral or nasal consonant. Coding of these vowels was initially done based on the judgements of a native-speaker linguist. Subsequently, F1, F2 and F3 were measured to validate the linguist's coding. Items exhibiting clear pronunciation errors were removed from the analysis (for example, a pronunciation that did not include a nasal coda for loanwords that had one, such as ['dʒɔpɪ] for the loanword *jump*). These errors were interpreted as being instances of misreading (in the Real Loanword Experiment) or misperception of the segment in coda position (in the Nonce Loanword Experiment). A total of 770 items were produced in the Real Loanword Experiment, and 600 items were produced in the Nonce Loanword Experiment. About 1% of the productions were errors (18 items in the Real Loanword Experiment, and 4 items in the Nonce Loanword Experiment).

Overall, participants used [ɐ] substantially more in the Real Loanword Experiment (82%, vs. 53% in the Nonce Loanword Experiment), both in oral context (75.6%, vs. 31% in the Nonce Loanword Experiment) and nasal context (88.4%, vs. 75% in the Nonce Loanword Experiment), which is consistent with the predictions presented above. In both experiments, the target vowel was predominantly produced as [ɐ] when followed by a nasal consonant. However, when the target vowel was followed by an oral consonant, it was produced as [ɐ] much more frequently when the item was a real loanword (e.g., *pub*, *bug*) than when it was a nonce loanword

² As per the definition in Davidson (2007), namely, the speakers who first acquire and introduce loanwords into their language.

(e.g., *vup*, *cush*). Figure 2 illustrates these results. In Figure 2, *other* corresponds to all non-[*ɐ*] productions. The distribution of such productions will be discussed in section 4.2 below.

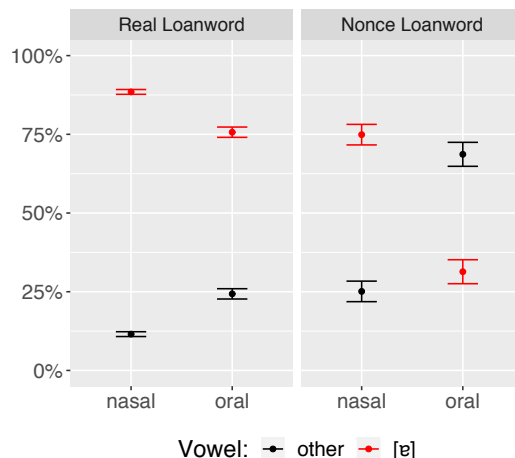


Figure 2: Proportion of [*ɐ*] by experiment and context.

These results are also consistent with the prediction that, in new loanwords, participants prefer [*ɐ*] in nasal contexts, and avoid it in oral contexts. These results appear to support the proposal that the ability to license [*ɐ*] in non-nasal contexts is part of the lexicon, but not part of the grammar. The lexicon and the grammar appear to be asymmetric when it comes to the adaptation of English /ʌ/: while speakers use the corresponding BP sound [*ɐ*] in loanwords that are already part of the lexicon, when acting like borrowers, they reject it in contexts where it cannot be licensed by the native grammar.

It is important to note that these results do not seem to be affected by the different methodologies employed in the two experiments. Since the Real Loanword Experiment used orthographic stimuli, it was possible that participants would observe the orthography of the English loanwords and prefer to use [*u*] instead of [*ɐ*]. This, however, was not the case. Even though the majority of *other* productions in the Real Loanword Experiment corresponded to [*u*], which reveals the influence of orthography to a certain extent, speakers used [*ɐ*] in the vast majority of their productions.

The data in Figure 2 were modelled with a hierarchical logistic regression in R (R Core Team, 2019). The model included *experiment* (real vs. nonce), *context* (oral vs. nasal) and *proficiency* (intermediate vs. advanced) as predictors, as well as an interaction between *experiment* and *context*, and by-speaker random slopes for *experiment* and *context*. The dependent variable was *vowel quality* ([*ɐ*] vs. other). Contrary to what was expected, the forms that participants produced for each of the target items in both experiments did not differ between the first and the second productions; for this reason, the model did not control for this variable.

Production of [*ɐ*] was significantly dispreferred in oral contexts, as well as in oral contexts in nonce loanwords. Participants' proficiency in English was not statistically significant, which suggests that knowledge of the target language does not impact the influence of the native grammar in loanword adaptation. Table 1 summarizes the results of the statistical model.

	Estimate	Std. error	z value	p value
intercept (real loanword:nasal context)	2.17	0.29	7.34	0.0001
nonce loanword	-0.68	0.51	-1.34	0.17
oral context	-1.02	0.33	-3.03	0.002
intermediate proficiency	0.12	0.25	0.47	0.63
nonce loanword:oral context	-1.62	0.33	-4.83	0.0001

Table 1: Statistical model for [*ɐ*] vs. *other* productions.

The next subsection discusses the properties of the non-[*ɐ*] outputs obtained in both the Real Loanword Experiment and the Nonce Loanword Experiment.

4.2 Non-[v] productions The vowels that were categorized as *other* had notably different qualities in the two experiments. In the Real Loanword Experiment, the majority of *other* productions corresponded to [u]. As mentioned above, this is not surprising given Portuguese orthography. Additionally, as previously illustrated, loanwords with /ʌ/ are optionally produced with [u] in the language (e.g., *bug* can be either [ˈbɛɡɪ] or [ˈbugɪ]).

In this experiment, 18% of the total productions were characterized as *other*. The quality of the produced *other* vowel varied depending on whether the context was oral or nasal. In oral context, 105 of a total of 108 *other* productions displayed the vowel [u] (the other three productions exhibited [ɔ]). In nasal context, 26 of a total of 36 *other* productions exhibited the diphthong [aʊ], all of which corresponded to the loanword *country*. Additional productions had one of the following vowels: [u] (4 productions), [ɛ] (3), [a] (2), and [o] (1). These numbers suggest that participants' preference for [v] in nasal contexts is stronger even when they are presented with written stimuli. Given written stimuli, participants might avoid [v] both in nasal and oral contexts, and prefer productions that are more in line with the loanwords' orthography. Although [v] was not avoided in neither context in this experiment, the fact that participants repair it more frequently in oral than nasal contexts when provided with orthographic stimuli further reflects the grammar's avoidance of [v] in non-nasal contexts.

In the Nonce Loanword Experiment, *other* productions were more varied, although, similarly to the Real Loanword Experiment, participants' choices were different depending on whether the target vowel was in oral or nasal context. Overall, 47% of all productions were classified as *other*. In nasal context, 45 of a total of 75 *other* productions exhibit [o], and 22 exhibit [u]. Other productions are [a] (4) and [ɛ] (4). [a] and [ɛ] are not possible in nasal context in BP,³ which suggests that these productions result from misperception of the novel loanwords. We turn to the use of [u] and [o] below. In oral context, 105 of a total of 205 *other* productions exhibit [a], 60 exhibit [ɔ], 34 display [u], 4 correspond to [ɛ] and 2 exhibit [aʊ]. Here, [ɛ] and [aʊ] should be the result of misperception, given that neither is acoustically similar to /ʌ/ nor corresponds to an allophone of /ʌ/ in BP.

Four possible productions deserve greater attention at this point: [u] (observed both in oral and nasal context), [o] (in nasal context), [ɔ] and [a] (both in oral context). The use of [u] in both contexts can be explained based on participants' knowledge of sound-letter correspondence patterns in English. Such a result was not initially predicted, as participants were instructed to produce the items based on what they heard.

On the other hand, although the use of [o] and [ɔ] cannot be explained based on orthography, it can be accounted for by a single explanation. English /ʌ/ is acoustically similar to BP [ɔ], so participants' production of [ɔ] in oral contexts can be explained based on the phonetic similarities between the native and the foreign segments. In BP, [ɔ] raises to [o] in nasal contexts (Wetzels, 1997; Barbosa & Albano, 2004). For this reason, loanwords perceived as having an [ɔ] in nasal context should be produced with an upper mid vowel in order to be adapted to the native system. It should be noted that, although the production of [ɔ] seems to be driven primarily by perception, it may be affected by the phonological system of the borrowing language as well. It is possible that participants perceive a sound similar to [v] in oral context in these loanwords, but, given the phonotactic constraints of their native grammar, avoid it in favor of a native phoneme that is also acoustically similar to /ʌ/. The precise role of perception and phonology in this particular adaptation strategy is a matter for future investigation.

Turning to [a], its use in oral contexts can be accounted for based on speakers' avoidance of [v] in native contexts. Since participants were instructed to produce the nonce loanwords using carrier sentences in BP, they strived to adapt the loanwords according to the patterns of the borrowing language. Since [v] and [a] are articulatorily very distinct, it is unlikely that participants perceived the target vowel as [a] instead of [v] (or /ʌ/; see the next section for discussion on what the input might be). Thus, it seems that BP's native phonology motivates participants' adaptation pattern where /ʌ/ results in [a].

Figure 3 illustrates the acoustic profile of the three most frequent productions in oral context in the Nonce Loanword Experiment. Vowel [u] is not included in the figure since I assume its production is motivated by orthography, not perception or phonology.

The results discussed in this subsection indicate that both perception and phonology strongly condition the adaptation of English /ʌ/ in BP: in oral context, perception is what prompts productions with [v], while phonology is what underlies productions with [a]. Productions with [ɔ] in oral context and [o] in nasal context

³ Certain BP varieties admit [ɛ] in nasal context, but as an allophone of /e/, not /a/ (Cagliari, 1977).

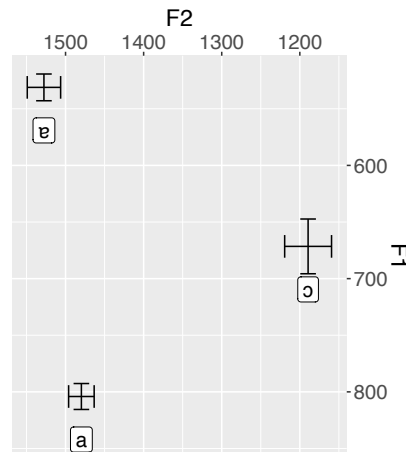


Figure 3: Vowels produced in oral context in the Nonce Loanword Experiment.

also seem to be triggered by perception, given the similarities between BP [ɔ] and English [ʌ]. However, as mentioned above, especially in oral context, it could also be the case that participants perceive /ʌ/ as [ɔ] because they do not expect to find [ɐ] in non-nasal context in native words.

The next section formalizes participants' productions in oral contexts in the Nonce Loanword Experiment using a probabilistic model of grammar.

5 Formalization

As discussed above, BP [ɐ] is only allowed in nasal contexts in native words. Thus, it is a mismatch between the source language and the borrowing language that prompts the use of different adaptation strategies in oral contexts. Even if we focus only on the productions that are not conditioned by orthography or due to misperception, we see that speakers use [ɐ] as well as [ɔ] and [a]. Both [ɐ] and [ɔ] seem to be fundamentally motivated by perception: as shown above, [ɐ] is the closest vowel to /ʌ/ in BP; BP [ɔ] is articulatorily very similar to English /ʌ/ as well, the crucial difference being roundedness.

The most frequent adaptation strategy in oral context among nonce loanwords involves [a]. This strategy is consistent with the distribution of [ɐ] and [a] in the BP native inventory, which suggests that it is affected by the phonological grammar of the language. What the use of these forms indicates is that, in the adaptation of English /ʌ/ in BP, constraints requiring faithfulness to the source must be in conflict with constraints requiring a ban on non-native patterns.

In this section, participants' outputs are formalized in a MaxEnt Grammar (Goldwater & Johnson, 2003; Hayes & Wilson, 2008), which allows for the probabilistic assessment of candidate forms. In MaxEnt, each constraint is weighted, and the number of violations that a given candidate incurs to a specific constraint is multiplied by the weight of such a constraint. Candidates are then assigned scores, which correspond to the sum of all violations multiplied by the respective constraint weights. Candidates that violate constraints with a higher weight are more severely penalized and are thus less likely to arise. MaxEnt values and probabilities for all candidates are calculated based on the scores of the output candidates. A framework such as MaxEnt is thus optimal for capturing non-categorical phenomena. Although MaxEnt tools are able to simulate constraint weights (see Hayes & Wilson (2008)), the present analysis employs weights that reflect the proportions and statistical results described in the previous section.

Before proceeding with the analysis, a few considerations about the nature of the input are necessary. Since participants were presented with auditory stimuli in the Nonce Loanword Experiment, I assume that the input corresponds to the forms that they were exposed to. In other words, the input does not correspond to an abstract representation of the novel loanwords, but instead is equivalent to the acoustic signal that the

participants heard. Thus, for a nonce loanword such as *vup*, the input would be [vʌp]. This assumption therefore entails that loanwords are first experienced by speakers through perception, at least when presented exclusively through auditory means (see Peperkamp & Dupoux (2003); Peperkamp (2004); Peperkamp et al. (2008); Davidson (2007)).

Given the three possibilities for the adaptation of /ʌ/ in oral contexts discussed in the beginning of this section, the constraints listed in (3) seem to be at play. Since outputs are expected to display an epenthetic vowel (as the examples in (1) illustrate), this analysis does not include any constraints related to syllable repair.

(3) FAITH: Every segment in the input is featurally identical to every segment in the output

FAITH(round): Every segment in the input is identical in [round] to every segment in the output

LICENSE-[e]: [e] is followed by a nasal consonant

Productions with [a] are the most frequent (42%), followed by productions with [e] (35%) and productions with [ɔ] (23%). Given the nonce loanword *vup*, candidate ['vɔpi] is the least likely to arise, followed by ['vɛpi] and ['vapi]. FAITH is the only constraint violated by ['vapi], as the candidate respects the native grammar's requirement with regard to the distribution of [e] (thus not violating LICENSE-[e]), and is faithful in [round] to the input (thus not violating FAITH(round)).

Candidate ['vɔpi] violates two of the three constraints presented in (3): FAITH and FAITH(round). Both of these violations are due to divergences with the input, one of which involves a change in the value of the feature [round]. Candidate ['vɛpi], on the other hand, violates only LICENSE-[e], as [e] arises in a context that is not licensed by the native grammar. This candidate, however, does not violate FAITH: since BP [e] and English /ʌ/ are acoustically and articulatorily very similar, and since BP does not have /ʌ/ in its inventory, speakers should not build a distinct representation for /ʌ/ when confronted with it. In a model that assumes perception as the first step in loanword adaptation, [e] thus corresponds to the faithful perception of English /ʌ/ by native BP speakers.

A comparison between the proportions of [a] and [e] suggests that the weight of LICENSE-[e] must be greater than the weight of FAITH; see the tableau in (4). The fact that outputs with [ɔ] are the least frequent indicates that the combined weights of FAITH and FAITH(round) must be greater than the weight of LICENSE-[e]. In the tableau, the weight of each constraint is presented in square brackets. The probabilities were calculated first by negating the score of each candidate and obtaining its exponential value, and subsequently by dividing the resulting exponential value by the sum of the exponential values of all the candidates. As can be seen in the tableau, such probabilities are very similar to the proportions observed in participants' productions.

(4)

vʌp	FAITH [0.5]	FAITH(rnd) [0.6]	LIC-[e] [0.7]	actual proportion	probability
'vɛpi			1	35.8	0.35
'vɔpi	1	1		23.2	0.23
'vapi	1			40.9	0.42

The constraint weights above, as well as the assumptions about the nature of the input, are not applicable to the loanwords that are already part of the BP lexicon. With respect to the input, speakers have abstract representations for the forms that are present in the lexicon, and such representations may have been construed based on the orthographic form of the loanwords, which would explain why the production of real loanwords with /ʌ/ often favors [u]. Orthographic influence aside, we can presume that candidates corresponding to inputs that stem from perception are subject to the same constraints as novel loanwords. In the case of real loanwords, however, outputs with [e] are the norm, and [a] is not observed. This suggests that, in the lexicon, the weight of FAITH must be overwhelmingly higher than the weight of constraints that require compliance with the native phonotactic patterns, such as LICENSE-[e].

6 Conclusion

In this paper, I examined the patterns of adaptation of English loanwords with /ʌ/ in Brazilian Portuguese. The closest segment in native BP inventory ([ɐ]) is used in the adaptation of such loanwords that are already part of the lexicon. However, as the use of [ɐ] is highly constrained in the native phonology (it is an allophone for /a/ allowed only in nasal contexts), it is unclear whether speakers generalize the non-native pattern, or whether they comply with the native phonotactic constraints, when they are confronted with novel loanwords.

To probe this issue, I conducted an experiment where participants produced real loanwords with /ʌ/, and another one where they produced nonce loanwords. The results showed that, in oral contexts, [ɐ] was used significantly more frequently in real loanwords than in nonce loanwords. Participants generally avoided [ɐ] in nonce loanwords where the target vowel is not followed by a nasal coda, preferring [a] instead.

The use of [a] in this context revealed that participants were using the native phonological grammar to filter the novel loanwords: in cases where they encountered [ɐ] in a context that is not licensed by the native grammar, they often used the corresponding licensed form. In a probabilistic constraint-based analysis, these patterns suggest that the weight of a constraint requiring compliance with the native phonological system (LICENSE-[ɐ]) must be higher than the weight of a constraint requiring faithfulness to the auditory input (FAITH).

Finally, these results indicate that native speakers avoid the generalization of non-native patterns that are present in the lexicon, mirroring what has been observed for the generalization of unnatural patterns in native grammars. It is puzzling, however, that non-native patterns are allowed in the lexicon but avoided in new loanwords, since one could expect that new loanwords would be more faithful to the source while those in the lexicon would have adapted to the native phonotactic patterns. It is possible that the loanwords that are part of the BP lexicon were introduced by bilingual individuals who have access to both BP and English phonological grammars (LaCharité & Paradis, 2005). In this case, loanword phonology is based on disseminators' representations for the source segment, which has been faithfully adapted. Another possibility is that recipients of the new loanwords (i.e., individuals who encounter loanwords after they have been adapted by a disseminator; Davidson (2007)) identify these items as foreign and strive to pattern them with loanwords that are already part of the lexicon. Further research is required to arbitrate between these possibilities, as well as to evaluate whether novel loanwords will eventually pattern with those that are already part of the lexicon.

References

- Barbosa, Plínio & Eleonora C. Albano (2004). Brazilian Portuguese. *Journal of the International Phonetic Association (illustrations of the IPA)* 34:2, 227–232.
- Battisti, Elisa (1997). *A nasalização no português brasileiro e a redução dos ditongos nasais átonos: Uma abordagem baseada em restrições [Nasalization in Brazilian Portuguese and the reduction of unstressed nasal diphthongs: A constraint-based approach]*. Ph.D. thesis, Pontifícia Universidade Católica do Rio Grande do Sul.
- Becker, Michael, Andrew Nevins & Jonathan Levine (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language* 88:2, 231–268.
- Cagliari, Luiz Carlos (1977). *An experimental study of nasality with particular reference to Brazilian Portuguese*. Ph.D. thesis, University of Edinburgh.
- Câmara Jr., Joaquim Mattoso (1970). *Estrutura da língua portuguesa [The structure of the Portuguese language]*. Petrópolis: Vozes.
- Davidson, Lisa (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology* 24, 261–286.
- Davis, Stuart & Mi-Hui Cho (2006). Phonetics versus phonology: English word final /s/ in Korean loanword phonology. *Lingua* 116, 1008–1023.
- Garcia, Guilherme D. (2017). *Weight effects on stress: Lexicon and grammar*. Ph.D. thesis, McGill University.
- Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. Spenader, Jennifer, Anders Eriksson & Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, Stockholm University, Stockholm, 111–120.
- Guimarães, Maximiliano & Andrew Nevins (2013). Probing the representation of nasal vowels in Brazilian Portuguese with language games. *Organon* 28, 155–178.
- Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:3, 379–440.
- Jarosz, Gaja (2017). Defying the stimulus: Acquisition of complex onsets in Polish. *Phonology* 34, 269–298.

- Kang, Yoonjung (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology* 20:2, 219–273.
- Kang, Yoonjung (2011). Loanword phonology. van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*, Wiley Blackwell, 2258–2282.
- LaCharité, Darlene & Carole Paradis (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36:2, 223–258.
- Leite, Yonne (1974). *Portuguese stress and related rules*. Ph.D. thesis, University of Texas at Austin.
- Massini-Cagliari, Gladis (1992). *Acento e ritmo [Stress and rhythm]*. Ed. Contexto.
- Paradis, Carole & Darlene LaCharité (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33, 379–430.
- Peperkamp, Sharon (2004). A psycholinguistic theory of loanword adaptations. *Proceedings of the Berkeley Linguistic Society* 30, 341–352.
- Peperkamp, Sharon & Emmanuel Dupoux (2003). Reinterpreting loanword adaptations: The role of perception. *Proceedings of the 15th International Congress of Phonetic Sciences*, 367–370.
- Peperkamp, Sharon, Inga Vendelin & S. Nakamura (2008). On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. *Phonology* 25, 129–164.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL <http://www.R-project.org/>.
- Silverman, Daniel (1992). Multiple scansion in loanword phonology: Evidence from Cantonese. *Phonology* 9, 289–328.
- Wetzels, W. Leo (1997). The lexical representation of nasality in Brazilian Portuguese. *Probus* 9:2, 203–232.
- Yip, Moira (2006). The symbiosis between perception and grammar in loanword phonology. *Lingua* 116, 950–975.